Order and disorder in crystals of hexameric NTPases from dsRNA bacteriophages

Article (Published Version)

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Order and disorder in crystals of hexameric NTPases from dsRNA bacteriophages

Erika J. Mancini,[a] Jonathan M. Grimes,[a] Robyn Malby,[a]† Geoffrey C. Sutton,[a] Denis E. Kainov,[b] Jarmo T. Juuti,[b] Eugene V. Makeyev,[b] Roman Tuma,[b] Dennis H. Bamford[b] and David I. Stuart[a]*

[a]Division of Structural Biology, The Henry Wellcome Building for Genomic Medicine, Oxford University, Roosevelt Drive, Oxford OX3 7BN, England, and [b]Institute of Biotechnology and Department of Biosciences, Vikki Biocenter, PO Box 56 (Viikinkari 5), 00014 University of Helsinki, Finland

† Present address: CSIRO, 343 Royal Parade, Parkville Vic 3052, Australia.

Correspondence e-mail: dave@strubi.ox.ac.uk

Received 4 June 2003
Accepted 22 August 2003

The packaging of genomic RNA in members of the *Cystoviridae* is performed by P4, a hexameric protein with NTPase activity. Across family members such as Φ6, Φ8 and Φ13, the P4 proteins show low levels of sequence identity, but presumably have similar atomic structures. Initial structure-determination efforts for P4 from Φ6 and Φ8 were hampered by difficulties in obtaining crystals that gave ordered diffraction. Diffraction from crystals of full-length P4 showed a variety of disorder and anisotropy. Subsequently, crystals of Φ13 P4 were obtained which yielded well ordered diffraction to 1.7 Å. Comparison of the packing arrangements of P4 hexamers in different crystal forms and analysis of the disorder provides insights into the flexibility of this family of proteins, which might be an integral part of their biological function.

1. Introduction

Many viruses couple nucleoside triphosphate (NTP) hydrolysis to nucleic acid translocation to package their genome into preformed viral particles. The structural basis of the NTP-dependent translocation of oligonucleotides has been extensively studied in a number of systems, in particular the portal complex of double-stranded DNA (dsDNA) bacteriophages (Baumann & Black, 2003; Catalano, 2000; Moore & Prevelige, 2002; Simpson *et al.*, 2000). The complexity of these systems, however, still obscures the basic mechanism of coupling. Members of the family *Cystoviridae* make attractive targets for studying the problem as the portal complex is composed of a single protein species, P4, which hydrolyses NTPs to provide the power for genome translocation (Juuti *et al.*, 1998; Kainov *et al.*, 2003). Icosahedral and single-particle cryoEM reconstructions have demonstrated that protein P4 from Φ6 is a ring-shaped hexamer positioned on the surface of the viral capsid at each fivefold symmetry axis. Such a symmetry mismatch, reminiscent of that observed at the portal vertex of dsDNA phages, could play a role in RNA packaging (de Haas *et al.*, 1999; Hendrix, 1998).

Much work has been devoted to the production of recombinant P4 from members of the *Cystoviridae* family (Φ6, Φ8 and Φ13) in order to characterize their *in vitro* activity (Kainov *et al.*, 2003). Recombinant P4 proteins have similar molecular weights of about 35 kDa and share a common ring-like hexameric architecture. However, they show very little sequence similarity (Kainov *et al.*, 2003). Furthermore, the NTPase activity of Φ8 P4 is strongly stimulated by RNA, whilst Φ6 and Φ13 are only weakly stimulated. Finally and most remarkably, isolated P4 from Φ8 and Φ13 possess a detectable *in vitro* RNA-translocating activity, which P4 from Φ6 lacks. Despite very low levels of sequence identity, a close inspection of the sequences suggests that P4 proteins from different viruses are likely to share significant structural similarity. Specifically, the presence of Walker motifs A and B and loosely conserved hexameric helicase motifs suggest the presence of a RecA-like fold and ATP-binding site. Atomic structures have now been reported for two hexameric helicases, but the mechanism for transducing the chemical energy generated by the NTP hydrolysis into oligonucleotide translocation remains elusive (Niedenzu *et al.*, 2001; Singleton *et al.*, 2000). The P4 proteins of the *Cystoviridae* are therefore likely to share a common core fold reminiscent of other hexameric helicases, but will undoubtedly show major rearrangements in the regions outside the conserved motifs. To verify this hypothesis and to help elucidate the mechanism of the mechano-chemical coupling of NTP hydrolysis with nucleic acid translocation, we have attempted to determine the atomic structure of these proteins by X-ray crystallography. The frustrating road to well ordered crystals is described in this paper.

2. Materials and methods

2.1. Protein expression and purification

Recombinant full-length Φ6 P4, clone Φ6 P4Δ310 (310/331 C-terminal truncation), full-length Φ8 P4 and full-length Φ13 P4 were expressed and purified as previously described (Kainov *et al.*, 2003; Ojala *et al.*, 1993).

A plasmid pDK10 was constructed to produce a C-terminal deletion mutant Φ8 P4Δ281 (Φ8 P4 missing amino acids 281–321). The 3′-terminal part of the P4 gene was PCR-amplified from pSJ1b encoding full-length protein (Ojala *et al.*, 1993) with recombinant Pfu DNA polymerase (Stratagene) using the primers 5′-TCGTCAACATATGGCTAGAAAAACGAAAGT-3′ and 5′-TGCGTAAGCTTAACCCCGCGAGACCA-3′ and inserted into pT7-7 plasmid (Tabor & Richardson, 1990) at *Nde*I–*Hin*dIII sites. Φ8 P4Δ281 was produced in the *Escherichia coli* B834 (DE3) strain transformed with pDK10 as described for Φ8 P4 (Kainov *et al.*, 2003).

2.2. Crystallization and X-ray analysis

All crystals were grown by the sitting-drop vapour-diffusion method. Unless otherwise specified, crystals were cryoprotected by transferring them into reservoir solution mixed with glycerol to a final concentration of 25% seconds prior to freezing in a nitrogen-gas stream. All data

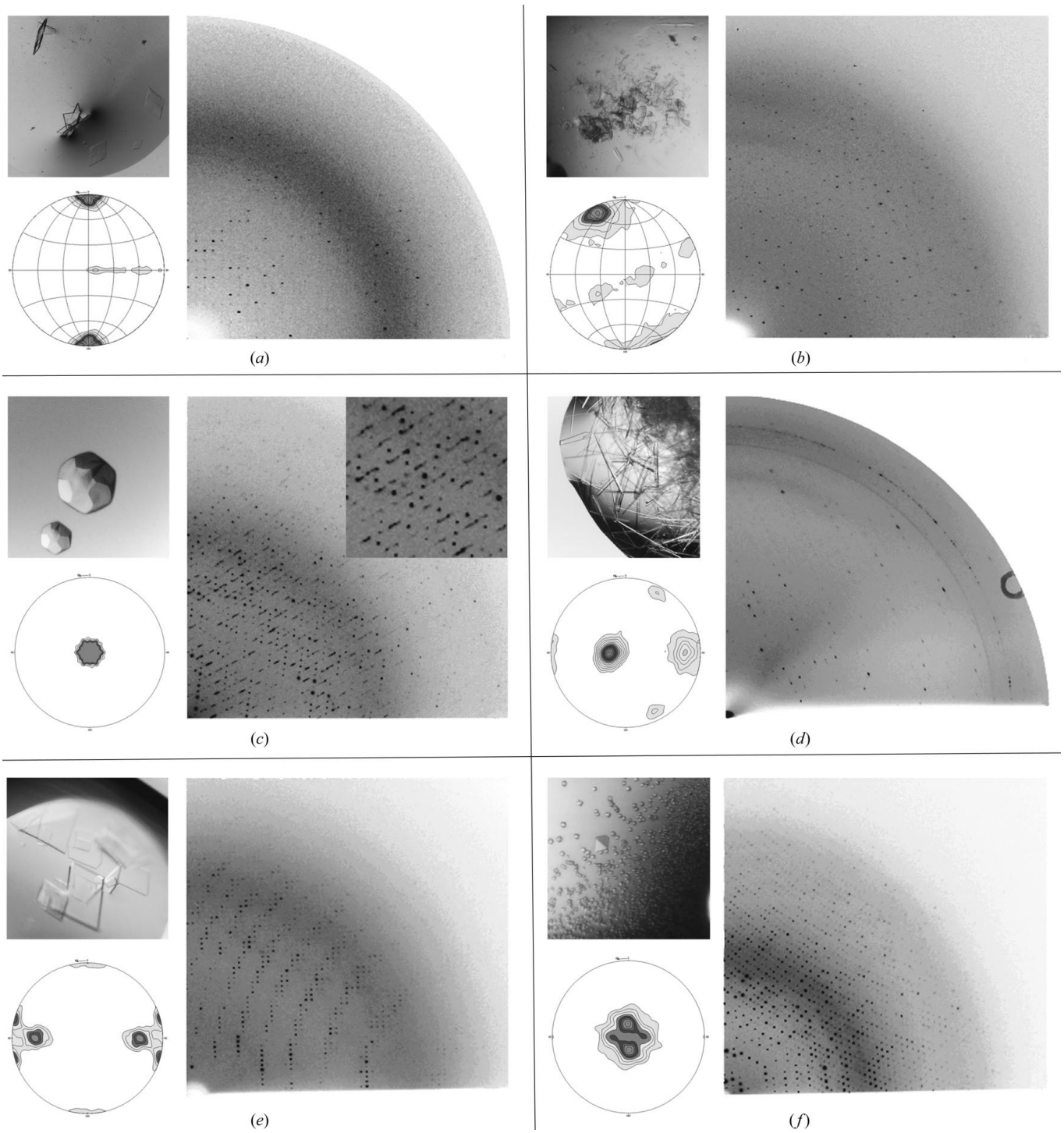


Figure 1
Crystals, diffraction patterns and self-rotation functions for different cystoviral P4s. (*a*) Φ6 P4Δ310, space group $P2_1$. (*b*) Φ6 P4Δ310, space group $P1$. (*c*) Φ8 P4, space group $P6_322$. (*d*) Φ8 P4Δ281, space group $C2$. (*e*) Φ8 P4Δ281, space group $P2_12_12$. (*f*) Φ13 P4, space group $P2_12_12_1$. The $\kappa = 60^\circ$ sections of the self-rotation functions calculated using the program *CNS* (Brünger *et al.*, 1998) are shown.

Table 1
Data-collection statistics.

Values in parentheses refer to the highest resolution shells.

	Φ6 P4Δ310	Φ6 P4Δ310	Φ8 P4	Φ8 P4Δ281	Φ8 P4Δ281	Φ13 P4
X-ray source	ESRF BM14	SRS 14.2	ESRF ID29	ESRF ID14-EH2	ESRF ID14-EH4	ESRF ID14-EH2
Detector	MAR CCD	ADSC Q4R CCD	ADSC Q210 CCD	MAR CCD	ADSC Q4R CCD	ADSC Q4R CCD
Wavelength (Å)	0.979	0.870	0.977	0.933	1.006	0.933
Space group	$P2_1$	$P1$	$P6_322$	$C2$	$P2_12_12$	$P2_12_12_1$
Unit-cell parameters						
a (Å)	98.9	93.8	100.7	168.1	139.0	97.0
b (Å)	102.5	95.2	100.7	111.4	159.6	116.0
c (Å)	171.3	98.3	131.2	103.0	100.4	149.1
α (°)	90.0	82.4	90.0	90.0	90.0	90.0
β (°)	90.1	86.2	90.0	106.8	90.0	90.0
γ (°)	90.0	64.9	120.0	90.0	90.0	90.0
Resolution range (Å)	20.0–3.5	20.0–3.0	30.0–2.0	60.0–3.4	20.0–2.9	20–1.7
Observations	181563	319853	1020966	186386	1254134	3208823
Unique observations	38948	72294	27264	26995	61037	180149
Completeness (%)	94 (92)	96 (90)	100 (100)	98 (94)	100 (100)	100 (98)
$I/\sigma(I)$	5.6 (2.5)	4.6 (1.1)	41.2 (14)	8.2 (1.7)	21.1 (6.0)	16.2 (2.4)
$R_{\rm merge}$† (%)	12.8 (48.5)	19.8 (62.0)	8.1 (38.0)	16.6 (44.4)	13.3 (51.2)	8.3 (54.9)
No. of hexamers in AU	2	2	1/6	1	1	1

† $R_{\rm merge} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$.

were processed and scaled using the *HKL*2000 suite of programs (Otwinowski & Minor, 1997). Statistics for data sets for the different crystal forms of the various proteins are summarized in Table 1.

Two crystals forms of Φ6 P4Δ310 were grown at 298 K from a 3.5 mg ml^{-1} protein solution in 20 m*M* HEPES–KOH pH 8.0, 5 m*M* $MgCl_2$, 2 m*M* $CaCl_2$, 5 m*M* ADP, 100 m*M* NaCl. The two forms belong to space groups $P1$ and $P2_1$, but were morphologically identical and appeared after nine months in a drop in which 3 µl of protein were mixed with 3 µl of a reservoir solution consisting of 6% PEG 4000 and 90 m*M* sodium acetate pH 4.5. The diffraction data for the $P2_1$ crystal form of Φ6 P4Δ310 were collected from three crystals at room temperature.

Crystals of full-length SeMet-derivatized Φ8 P4 were grown at 298 K using 0.1 *M* sodium acetate pH 4.6 and 2.2 *M* ammonium sulfate as a precipitant. Drops consisted of 0.9 µl of protein at a concentration of 3 mg ml^{-1}, 0.9 µl of reservoir solution and 0.4 µl of 100 m*M* DTT. Crystals of the truncation mutant Φ8 P4Δ281 at a concentration of 5 mg ml^{-1} were grown at 298 K using a Cartesian crystallization robot (Walter *et al.*, 2003) with a drop size of 100 nl of protein and 100 nl of precipitant (0.1 *M* Tris pH 8.0, 18% PEG 1000). A second crystal morphology for Φ8 P4Δ281 was found using the crystallization condition for the full-length protein, six months after setting up the crystallization drops.

Crystals of full-length Φ13 P4 at a concentration of 12 mg ml^{-1} were grown at 293 K using 0.1 *M* Tris–HCl pH 7.0, 1 *M* trisodium citrate and 0.2 *M* NaCl as a precipitant. The dropsize was, again, 100 nl.

3. Results and discussion

3.1. Crystal-packing arrangements of P4 from Φ6

Several crystals of different morphology were rapidly obtained from purified active full-length Φ6 P4, but none diffracted to high resolution. C-terminally truncated proteins comprising at least 290 out of 331 amino acids retain NTPase activity. Clone P4Δ310, which is truncated at residue 310, delivered several crystals in a single drop ten months after setting of the trays. The first crystals from that drop diffracted to 3.5 Å and belonged to space group $P2_1$. A second crystal picked up from the same drop a month later diffracted to 3.0 Å and belonged to space group $P1$. Analysis of their self-rotations and native Patterson functions revealed the presence of two hexamers in the asymmetric unit of both space groups in the same orientations and related by a simple translation (Figs. 1 and 2). The two hexamers, *A* and *B*, which are non-crystallographically related in space group $P2_1$ are crystallographically related in space group $P1$; conversely, the crystallographically related molecules in space group $P2_1$ are non-crystallographically related in $P1$ (Fig. 2). It seems plausible, given the related arrangement, that dehydration of the drop caused the transformation of the space group from $P2_1$ to $P1$. This suggestion is in line with a drop in solvent from ∼40%, calculated for space group $P2_1$, to ∼34% for space group $P1$ and the improved resolution of the diffraction (Esnouf *et al.*, 1998).

3.2. Crystal-packing arrangements and disorder of P4 from Φ8

Crystals of full-length P4 Φ8 belong to the apparent space group $P6_322$, with unit-cell parameters $a = b = 100.7$, $c = 131.2$ Å, $\alpha = \beta = 90.0$, $\gamma = 120.0°$. The unit-cell dimensions are consistent with the presence of one monomer of P4 in the crystallographic asymmetric unit. These crystals, however, gave rise to a rather complex diffraction pattern (Fig. 1*c*) that included a honeycomb-like effect and diffuse rods and bands in addition to the Bragg reflections. Such effects are characteristic of statistically disordered crystals, first reported in 1954 for imidazole horse methaemoglogin (Bragg & Howells, 1954; Cochran & Howells, 1954). This is an unfortunate and intimate type of twinning in which crystal domains showing systematically different orientations or positions are distributed randomly throughout the crystal to produce the characteristic blurring of the Fourier transformation in a systematic way.

The diffraction of Φ8 P4 can be explained by regarding the crystals as being based upon a perfect lattice of hexamers organized in a smaller unit cell ($a = b = 100.7$, $c = 65.6$ Å). The sheets of hexamers in the c planes have a well defined structure and are capable of yielding sharp reflections. However, successive sheets are stacked together with random fractional displacements of 1/2 in the a or b direction. The relative displacement of the layers manifests itself as an apparent doubling of the unit-cell dimension in the c direction. The reflections of this superlayer are diffuse, leading to an interesting pattern containing sharp lines and diffuse lines of various widths (Fig. 1*c*).

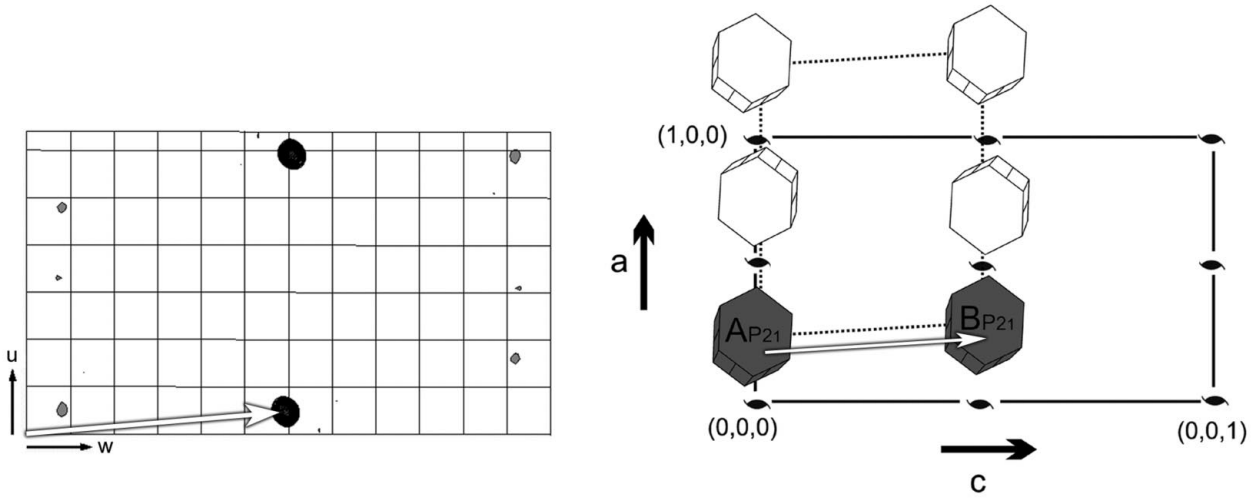
As for Φ6 P4, in Φ8 P4 the C-terminus is not required for enzymatic activity. Clone Φ8 P4Δ281, truncated at residue 281 out of 321, was soluble, expressed at high levels in *E. coli* and crystallized to yield weak diffraction to 3.4 Å (Fig. 1*d*). Also inline with Φ6 P4Δ310, crystals of Φ8 P4Δ281 were found in drops several months after setting up the trays. These crystals belonged to space group $P2_12_12$ and displayed ordered diffraction to better than 2.9 Å resolution (Fig. 1*e*). The improved diffraction, as in the case of Φ6 P4, may be a consequence of proteolytic cleavage, dehydration or a combination of the two.

3.3. X-ray analysis and packing arrangements of P4 from Φ13

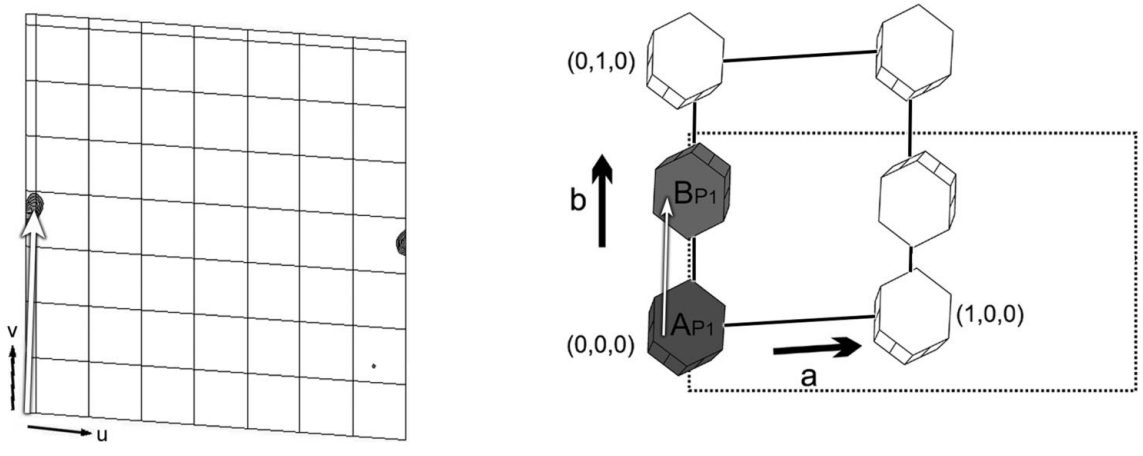
Crystals of full-length P4 from Φ13 diffracted strongly to better than 1.7 Å resolution (Fig. 1*f*). The unit-cell parameters are consistent with the presence of one hexamer of Φ13 P4 in the crystallographic asymmetric unit, assuming a solvent content of 33% (Table 1). The very low solvent content accords with the strong diffraction from these crystals. The self-rotation function shows a strong peak at $\psi = 74.0$, $\varphi = 90.0$, $\kappa = 60.0^\circ$ consistent with a non-crystallographic sixfold axis canted about 15° from the *c* axis (Fig. 1*f*).

3.4. Packing arrangements and their biological significance

The surface of the P4 hexamer, possibly owing to its biological role as a molecular interface, seems to offer an incredibly good interface for crystallization in a variety of different space groups (six different space groups are reported in Table 1). However, this ability to form varied interactions appears to bring with it a tendency towards stacking disorders and changes in crystal lattice over time (perhaps in response to dehydration). Obtaining ordered diffraction was a bottleneck in our structural studies of P4. In this case, clues from anisotropically diffracting crystals were used to engineer truncated constructs with somewhat improved crystal packing; however, species variation was ultimately the key to obtaining good diffraction reliably. Clearly, a genomic approach in which the proteins are examined from several related organisms of the same family is likely to be an efficient general strategy for obtaining diffraction-



(*a*)



(*b*)

Figure 2
Native Patterson functions and crystal-packing arrangements of the hexamers in related space groups $P2_1$ and $P1$ of Φ6 P4Δ310. Native Patterson functions were calculated in *CCP*4 (Collaborative Computational Project, Number 4, 1994) and the figures were drawn with *GROPAT* (R. Esnouf, unpublished program). (*a*) The native Patterson function for the $P2_1$ crystal form, shown on the left and calculated using data between 20 and 3.5 Å, produced a 13.5σ peak at fractional coordinates $a = 0.078$, $b = 0.391$, $c = 0.492$ originating from NCS-related molecules. The packing of hexamers for space group $P2_1$ is shown on the right, with NCS-related molecules *A* and *B* coloured grey. (*b*) The native Patterson function for the *P*1 crystal form, shown on the left and calculated using data between 20 and 3.5 A, produced a 8.1σ peak at fractional coordinates $a = 0.986$, $b = 0.472$, $c = 0.5$ originating from the NCS-related molecules. The packing of hexamers for space group *P*1 is shown on the right with NCS-related molecules *A* and *B* coloured grey.

quality crystals, as pioneered nearly 50 years ago by John Kendrew.

This work was supported by the Human Frontier Science Programme, the Academy of Finland [Finnish Centre of Excellence Program 2000–2005, grants 172623 (RT), 1202855 and 1202108(DHB)] and the EC (SPINE, QoL-2001-CE18). EJM was the recipient of an EMBO long-term fellowship, DEK is a fellow of the National Graduate School in Informational and Structural References, JMG is supported by the by the Royal Society and DIS by the UK Medical Research Council.

References

Baumann, R. G. & Black, L. W. (2003). *J. Biol. Chem.* **278**, 4618–4627.

Bragg, W. L. & Howells, E. R. (1954). *Acta Cryst.* **7**, 409–411.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Catalano, C. E. (2000). *Cell Mol. Life Sci.* **57**, 128–148.

Cochran, W. & Howells, E. R. (1954). *Acta Cryst.* **7**, 412–415.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst*, D**50**, 760–763.

Esnouf, R. M., Ren, J., Garman, E. F., Somers, D. O., Ross, C. K., Jones, E. Y., Stammers, D. K. & Stuart, D. I. (1998). *Acta Cryst.* D**54**, 938–953.

Haas, F. de, Paatero, A. O., Mindich, L., Bamford, D. H. & Fuller, S. D. (1999). *J. Mol. Biol.* **294**, 357–372.

Hendrix, R. W. (1998). *Cell*, **94**, 147–150.

Juuti, J. T., Bamford, D. H., Tuma, R. & Thomas, G. J. Jr (1998). *J. Mol. Biol.* **279**, 347–359.

Kainov, D. E., Pirttimaa, M. J., Tuma, R., Butcher, S. J., Thomas, G. J. J., Bamford, D. H. & Makeyev, E. V. (2003). Submitted.

Moore, S. D. & Prevelige, P. E. Jr (2002). *Curr. Biol.* **12**, R96–R98.

Niedenzu, T., Rèoleke, D., Bains, G., Scherzinger, E. & Saenger, W. (2001). *J. Mol. Biol.* **306**, 479–487.

Ojala, P. M., Juuti, J. T. & Bamford, D. H. (1993). *J. Virol.* **67**, 2879–2886.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J., Olson, N. H., Morais, M. C., Grimes, S., Anderson, D. L., Baker, T. S. & Rossmann, M. G. (2000). *Nature (London)*, **408**, 745–750.

Singleton, M. R., Sawaya, M. R., Ellenberger, T. & Wigley, D. B. (2000). *Cell*, **101**, 589–600.

Tabor, S. & Richardson, C. C. (1990). *J. Biol. Chem.* **265**, 8322–8328.

Walter, T. S., Diprose, J., Brown, J., Pickford, M., Owens, R. J., Stuart, D. I. & Harlos, K. (2003). *J. Appl. Cryst.* **36**, 308–314.